口语句子的韵律边界：窥探言语理解的秘窗

方岚[1,2] 郑苑仪[3,4] 金晗[2] 李晓庆[3,4] 杨玉芳[3,4] 王瑞明[1]*

([1]脑认知与教育科学教育部重点实验室，华南师范大学心理学院/心理应用研究中心, 广州 510631)

([2]广州大学外国语学院，广州 510030) ([3]中国科学院心理研究所行为科学重点实验室, 北京 100101)

([4]中国科学院大学心理学系, 北京 100049)

chinaXiv:202010.00070v1

摘要 韵律边界加工与言语理解紧密相关，最近十几年来逐渐成为心理学和语言学的研究焦点。韵律系统包含若干由小到大的韵律单位，不同单位的韵律成分其边界强度不同，表现在音高、延宕和停顿三个声学线索上的参数也不同。句子的听力理解过程中，听话人运用声学线索感知权重策略对韵律边界的声学线索进行加工。从神经层面上来看，对于韵律边界的加工，大脑显示出独立且特异性的神经机制。韵律边界的加工能力在婴儿出生后随年龄的增长而发展，到了老年阶段则逐渐退化，而且似乎能够对二语迁移。未来，需要扩大对韵律边界声学表现的考查范围，进一步明确韵律边界的加工过程，进一步厘清韵律边界加工和句法加工之间的关系，进一步关注二语者韵律边界加工能力的发展。

关键词：韵律边界；言语理解；声学表现；神经振荡；感知发展

1 引言

韵律是口语理解的核心（Frazier et al., 2006; Holzgrefe et al., 2013）。韵律边界，作为口语韵律特征的重要组成部分，是把语流切分成不同层次的言语组块的组织框架（Frazier et.al., 2006），其位置的设定和声学特征的表现都会影响听者对句子的理解。比如，它能够促进句子的理解，能够解决句法歧义，能够帮助听者分解连续的言语信号、提取意义单位，从而习得语言 （Holzgrefe-Lang et al., 2016； Wagner & Watson, 2010）。韵律边界的加工能力与阅读理解的流畅性也紧密相关（Holliman et al., 2017）。显然，要了解韵律和语言加工，韵律边界是不能错过的窗口。研究者围绕韵律边界开展了众多的研究，特别是最近十几年，新技术的发展带来了更多的新发现。

2 韵律边界的层级结构

严格分层假说（Strict layering hypothesis）认为韵律系统包含 6 个层次，音节（syllable）、

---

* 收稿日期: 2020-05-04

本文系广东省哲学社会科学“十三五”规划项目“中国 EFL 学习者韵律边界加工机制研究”(GD19WXZ11) 的阶段性成果。

通信作者: 王瑞明, E-mail: wruiming@163.com

音步（foot）、韵律词（clitic）、音系短语（phonological phrase）、语调短语（intonational phrase）和话语（utterance）（Selkirk, 2005）。在音系结构中，每一个韵律成分由比它低一级的韵律成分构成，每个韵律成分都有它的边界。同一层次的成分通过这个边界区分开来。这样，韵律边界和其它的韵律特征（比如重音）就构成了言语交际中词语切分和句法解析的框架。

韵律边界的层级在口头交际中能够被说话人和听话人区分，区分的依据是边界层级所对应的边界强度，而强度由边界声学线索的参数表征（Cole, 2015; Krivokapic, 2012）。这些线索是（1）音高变化，即通过音高上升、音高下降或者再设（下降后再上升回复到原有高位）表示边界音的存在（也称为 f0, 二者的区别在于前者从感知的角度，而后者从发音的角度进行描述)；（2）音长，即边界前词语延长的持续时间；（3）停顿，即两个韵律单位之间的间隔性中断 （Männel & Friederici, 2016）。

研究者对韵律层级与语法层级的关系进行了一定的研究，认为两者相关却又不是一一对应。有研究者认为，韵律层级和句法结构之间存在着映射关系，每个韵律层级都有相应的句法单位存在，韵律结构边界和句法结构边界常常一致，尤其是在朗读中。据此现象，Truckenbrodt （1995）提出了 Generalized Wrap 理论，认为语调短语边界倾向于不打破句法短语，即韵律边界与句法节点保持一致。但 Frazier et al.（2004） 的研究分别以语法节点和意思单元（sense unit）为条件来操控语调短语边界位置，以语句自然度判断为任务，发现句法节点不是韵律边界位置的绝对性限定因素。显然，韵律结构的复杂度远远小于句法结构，也使得韵律边界和句法边界之间的关系不可能是单纯的一对一关系。除此之外，日常经验显示，自然谈话中韵律边界和句法边界的关系更加松散。这类语言现象中又蕴含着怎样的规律，有待进一步发掘。

3 韵律边界的加工线索

在语音听辨的过程中，虽然多个声学线索同时呈现，但是听者根据声学线索所具有的分辨力而被赋予不同的权重，这一理论即言语感知线索权重理论 （Qin et al., 2016）。根据这一理论，言语感知是多维度的，不同的范畴、不同的语言赋予声学线索不同的权重，因而成为边界声学线索感知研究的重要理论。

成人母语者对韵律边界声学线索的利用是语言特异性的。比如，英语母语者在判断从句边界时，对界前音素延长的倚赖多于停顿（Aasland & Baum, 2003），而对于汉语母语者来说，停顿是最重要的边界线索（Yang et al., 2014）。德语母语者使用停顿的范畴和界前音素延时来判断边界，而对音高的倚重次之（Petrone et al., 2017）。值得一提的是，韵律边界感知研究

chinaXiv:202010.00070v1

采用的方法多是用某个线索的去留来判断其重要性，忽视了可能存在的声学阈值。即，在某个范围内，某个声学线索是有效的，范围之外则失去效力。Petrone et al.（2017）首次使用阈值的概念来探讨声学线索的权重，为后来的研究提供了一个新的方向：我们可以进一步从范畴性和连续性的角度深化边界线索知觉的研究。

关于韵律边界加工是基于局部线索还是整体线索也存在争议。一种观点认为（e.g., Marcus & Hindle, 1990） 韵律边界的加工是依据紧依着边界的韵律线索，无须其他的韵律信息。因而，韵律边界的加工也是局部的，独立于上下文的。上述关于边界感知权重线索的研究都基于该观点。另一种观点则认为，韵律边界的加工是整体性的，相邻韵律短语边界的强度会影响韵律边界的感知。例如，old men#1 and women # 2 with very large houses 这样的口头语句中，如果#1 处的边界强度大于#2 处的边界强度，那么该短语的意思是“老年男士们和拥有豪宅的女士们”，句中并没有包含老年男士们有没有豪宅的信息； 如果#2 处的边界强度大于#1 处的边界强度，那么其意思就变成了“拥有豪宅的老年男士们和(老年)女士们”，女士们也拥有豪宅（Clifton et al., 2002）。 也就是说，听话人需要在整体框架下加工韵律边界的线索。

研究者还对韵律边界加工和句法加工的关系进行了探讨。韵律边界是切分语流的组织框架，句法是语言组织的规则，二者都具备组织功能，因此相互影响。一方面，句法对语调边界感知有预测作用。Cole et al.（2010）根据 Buckeye 语料库的数据和 97 名被试对上述语料的转写结果，发现从句的句法边界是最强大的边界感知预测指标。Buxó-Lugo & Watson （2016） 进一步发现句法短语边界位置的期望也影响语调边界的感知：被试在句法许可位置上对韵律边界的感知效率要显著高于句法不许可位置的韵律边界。例如，在口头句子（1）Put the big#1 bowl on the tray 和（2）Put the bowl that’s big#2 on the tray 中，分别于 big 之后设置韵律边界。韵律边界#1 处为句法不许可处，#2 则为句法许可处。借助声学软件合成不同强度的韵律边界（改变 F0 的廓形、停顿的时长和界前元音的延长比率）。结果发现，被试对韵律边界判断的实验任务中，不论在哪个韵律边界强度等级下，被试在许可处报告韵律边界的比例总是要多于不许可处。该结果说明听话人对韵律边界的检测并不是严格地由听觉因素驱动，而是受句法因素影响的。结构明晰的句法也被证明能够促进韵律边界的感知（Honbolygo et al., 2016）。

另一方面，韵律边界影响口头句子的结构解析。Kjelgaard & Speer （1999）的实验使用花园路径句作为语言材料，发现恰当的韵律边界能够克服句法歧义。作者认为这是因

chinaXiv:202010.00070v1

为在语句加工的过程中，韵律边界充当了句子结构的指示器。这个研究似乎提示韵律边界是句法结构在语音层面上的表征。通过韵律，我们得以窥见深层次的句法结构信息。这个效应在 Steinhauer et al. （1999）的研究中再次得到 ERP 证据的证实。随后，边界删除假说（Boundary Deletion Hypothesis） 强化了韵律边界在句法解析中的作用：如果在一个本不该有韵律边界的地方增加了一个边界，那么所需要的时间比错过一个韵律边界（即，本该停顿却没有停顿）更多 （Pauker et al., 2011; Bögels et al., 2013）。这些研究似乎暗示在口语中韵律结构对句子结构解析的作用要先于句法引发的解析，在句子理解中并不是仅仅为辅助的角色。之后，Webman-Shafran & Fodor (2016)的研究发现，在加工双介词短语结构（pp-construction）的短句如 hid his anger about the divorce from Tami 时，如果第二个介词短语（from Tami）前有韵律边界，那么听话人倾向于将第二个介词短语作为 hid 的修饰语（即，high attachment 高附着）；如果没有，则将它当成是 the divorce 的修饰语（即，low attachment 低附着）。显然，这个结果挑战了语言学中的晚闭合原则。根据晚闭合原则，双介词结构中的第二个介词短语应该是低附着的(即修饰最近的单位)，但是，由于韵律边界的存在，第二个介词结构的附着就变得有两个可能：若第二个介词结构前不存在韵律边界，听话人采用低附着的句法结构解析方案；若存在一个韵律短语边界的时候，听话人则倾向于采用高附着的方案（Webman-Shafran & Fodor, 2016）。Webman-Shafran 和 Fodor（2016）的研究似乎更进一步凸显了韵律边界在句子理解中的作用——它着实决定了句子的结构和句意。这样的结果让我们理解韵律边界在句子加工中的角色变得更为复杂，但是也给句子加工模型带来了新的视角。

模块论（Module Approach）和交互论（Interactive Approach）分别是针对句子加工机制的两大理论流派。模块论的主要观点是听话人在加工句子时，首先根据语言学的句法建立句子的结构框架，然后再将词语分配到不同的句法位置上，最后提取语义合成句意。交互论的主要观点则认为对句法和非句法信息（例如语义、语用和韵律等）的应用是并行的。现有模型强调语义/语用的贡献，对韵律贡献的讨论非常少（两种理论总是交替成为相关领域的主流理论，其钟摆式历史可参见 Clifton & Duffy, 2001）。然而由上文所述可见，口语句子理解的加工过程中韵律边界通过韵律组块（prosodic chuncking）的方式参与句法解析（syntactic parsing）, 影响听话人对句子意思的理解。具体来说，听话人首先知觉声学特征然后投射到句子的结构上，完成韵律组块（chunking），产生句子结构的解析结果。然而，现有基于模块论的加工模型忽略了口语韵律在其中的作用。对于口语加工，韵律边界

chinaXiv:202010.00070v1

同样能够起到切分不同句法成分的作用，从而促进词汇在不同位置、句法角色的分配，所以应当考虑纳入韵律规则来更新模型。另一方面，对于交互论而言，韵律边界显然对语言学的句法解析有限制和引导作用，二者并行参与在线句子加工，其核心观点得到支持。但是，韵律和其他维度信息之间的交互和整合是如何开展的，值得深入探究。总的来说，传统的句子加工模型主要关注语言学句法和词汇意义在句子加工中的贡献，而关于韵律边界和句法交互关系的讨论将对拓展已有的口语句子加工方法论（关于模块轮和交互论的一些具体模型可参见 Clahsen & Felsh, 2017）有重要的启迪意义。

4 韵律边界加工的神经响应

最近十几年，研究者特别关注韵律边界加工的神经响应。从神经层面上探讨韵律边界是如何被人类听觉系统获取而用以促进语音感知的，以及其背后的认知神经机制是如何的。

4.1 CPS 与韵律边界加工

首先，研究者提出的关键问题是，韵律边界是一种独立的加工机制，还是口语加工过程中的“副产品”？换言之，对于韵律边界的加工是否存在特异性的认知神经机制呢？对此，研究者采用 ERP（Event-related Potential，事件相关电位）技术进行了回答。Steinhauer et al.（1999） 在 Nature 的子刊 Neuroscience 杂志上首次报告了韵律边界加工的特异性脑电成分 CPS（Closure Positive Shift，闭合正偏移成分）。该研究要求 20 名大学生被试在实验中完成判断韵律是否恰当的任务。结果发现在语调边界处存在特异性的脑电成分，这种成分潜伏期为 400-500ms，分布于头皮中后部。为了剥离语义和句法因素的影响，验证 CPS 的特异性，Pannekamp et al.（2005）使用 ERP 技术开展了实验，分别检测被试对四种信息量逐次递减的句子的加工。这四种句子分别是：1.正常句子，2.无意义句（jabberwocky sentence），句子中的实词被假词替代，无语义信息，3.假句 （pseudo sentence），句子中的实词和功能词被假词替代，无句法和语义信息，4.嗡鸣句（hummed sentence），整个句子由嗡嗡声组成，只有韵律信息。实验结果显示，在每一种句子的韵律短语边界，都观察到了 CPS 成分，包括只有嗡嗡声的句子。研究结果说明在句法、语义信息缺乏的情况下，仍会诱发 CPS 成分，即韵律边界的加工是独立于句法、语义和音素的加工单独存在的。Honbolygo et al.（2016）分别以嵌入句句型（如，The grandfather, who entered, was thirsty.）的真句和假句为刺激材料进一步验证了韵律边界的感知是独立存在的。CPS 还被证明具有跨语言的普遍性。例如，在英语 （e.g., Pauker et al., 2011; Steinhauer et al., 2010），荷兰语 （e.g., Bögels et al., 2010），

chinaXiv:202010.00070v1

中文 （Li & Yang, 2009） 和韩语中（Hwang & Steinhauer, 2011） 都被证实存在 CPS。

为了进一步探究 CPS 与不同层级韵律边界的关系，Li 和 Yang 采用了中文句子（2009）为材料，发现不同层级韵律边界均能够稳定地诱发 CPS，即使在去除了边界处停顿线索的情况下，CPS 仍然存在。并且，CPS 的潜伏期和韵律层级之间存在系统性变化，即随着韵律层级的增加（韵律短语、语调短语、小句边界、联句边界），CPS 潜伏期增加。这表明，人脑不仅能够独立地感知韵律边界，且能够对韵律边界的不同层级进行特异性表征，从而将长语音流切分、整合为多层级嵌套的语言结构。

4.2 大脑神经振荡活动与韵律边界加工

CPS 反映了与韵律边界处相关的、锁时锁相的神经反应，而大脑神经振荡活动（neural oscillation）则能够进一步揭示大脑锁时非锁相的节律性神经活动，更直观地反映大脑对听觉语音信号中韵律边界的感应活动。研究者采用 EEG 和 MEG 记录被试听语音时的脑电信号，发现听语音时，人类大脑听皮层的神经振荡活动与低频段（<10Hz）语音包络存在较强的同步节律变化（e.g. Brodbeck et al., 2018; Ding et al., 2016; Ding et al., 2018），这一频段对应于韵律边界划分的不同语言层级的节律，例如 delta 频段（0～3Hz）的调制包络与句子或短语节律有关，theta（4~10Hz） 的调制包络与音节节律有关（Poeppel, 2003）。研究者将这一机制形象地喻为大脑对语音信号产生的“夹带”作用（entrainment），认为大脑神经振荡活动如同“神经录音器”般在神经系统上还原语音的节律变化。基于这一机制，听觉系统得以获得韵律边界知觉，将输入的长语音流切分为小尺度的语音单元进行加工。“夹带”作用被证明能够优化语音理解过程，听有意义语音时“夹带”作用强于听无意义语音，并且低频段的“夹带”作用能够预测语音理解成绩（Gross et al., 2013; Park et al., 2015）。

尽管“夹带”作用与韵律边界感知密切相关，但其背后反映的加工机制仍是不明确的。一种观点认为“夹带”作用纯粹是声学包络引起的，而和语言信息的理解无必然的联系。他们发现人们在听不可理解的倒转语音时也能够产生“夹带”作用（Howard & Poeppel, 2010）。另一个观点则支持“夹带”和“语言理解”存在密切关系，认为“夹带”只在可懂语音条件下产生、或在高可懂度条件下“夹带”作用更强（e.g. Gross et al., 2013; Park et al., 2015）。例如，Doelling et al.（2014）采用语言（delta 节律呈现的数字序列，3Hz 左右）作为材料，发现去除与节律信息有关的波动线索后，大脑对 delta 频段（2～4Hz）的“夹带”活动显著降低，而加入能量波动线索后“夹带”得到了恢复。除此之外，这一“夹带”活动强度还与听者对语音尖锐度（sharpness）和可懂度（intelligibility）的感知评分存在正相关关系。针对

chinaXiv:202010.00070v1

该类“夹带”在言语理解中的机制，Giraud 和 Poeppel（2012）提出了基于神经振荡的层级绑定模型 （Oscillation-based Hierarchical Binding Model），认为语音单元左边界（edge）上的包络调频信号输入，能够引起听觉系统 delta-theta 频段（1～8Hz）神经振荡活动的相位重设（phase reset），进而调整 gamma 频段（30～70Hz）神经振荡活动波幅，使神经兴奋活动起伏与语音声学包络的起伏相对齐，为加工随后而来的语音信息提供了“时间框架”与“神经兴奋性准备”。在这一过程中，大脑听皮层实现对语音信号包络的同步“调协”产生“夹带”作用，从而得以将语音流信号信息进行切分与解码。

对声学包络信号的自下而上“解析”加工并非引起“夹带”作用的唯一原因，某些自上而下的过程被证明也能引起大脑对语音的“夹带”作用。例如，Ding et al.（2016）考察了去除韵律边界声学线索时是否仍然存在“夹带”作用。他们采用固定节奏的四字中文短句为材料（如：绵羊吃草），使得音节、词汇和句子节律分别为 4Hz、2Hz 和 1Hz，仅保留语义和句法线索，而去除韵律边界处停顿、边界前音节延长等声学线索。基于神经振荡活动的分析发现，在不同层级语音结构对应的节律上均产生了显著的“夹带”活动（4、2、1Hz）。Ding 和 Jin（2019）研究进一步表明“夹带”作用与基于句法或任务的语音组块规律有关，而并非加工词汇语义联系的副产物。在鸡尾酒会中，听者对“注意”语流的“夹带”效果更好（Brodbeck et al., 2018; Ding et al., 2018）。另一些研究证据则从声学线索的角度说明，除了语音包络上的调频节律信息，与韵律边界无关的、仅表征“音位”特征的精细频率信息也能够导致“夹带”作用的产生（Ding et al., 2014; Zoefel & VanRullen, 2015, 2016）。以上研究结果表明，自上而下的语音、语义和句法分析过程以及选择性注意都能够调控语音流切分的神经活动，从而导致“夹带”作用的产生。这些结果表明，大脑不同频段神经振荡活动的“夹带”作用，更可能是大脑基于自下而上加工与自上而下调控的交互过程进行语音流切分和重组的“混合物”（mixture），也即大脑对语音包络的“夹带”作用也包含某些“主动”的感知加工过程。

那么，“夹带”作用是否也反映“韵律边界”本身的自上而下加工过程呢？已有的一些研究证据表明，“夹带”作用与对韵律边界的时间预期的确存在相关。研究者将连续语音流分为语境 （context） 和目标短语 （target word），发现听者能够基于语境中韵律边界的划分规律，对之后的语音边界进行预测。例如，Dilley 和 Pitt （2010）采用功能词有无不影响通顺度的语句为材料（e.g. Deena doesn’t have any leisure or time…），将语境部分（Deena doesn’t have any lei…）节律放慢 （时长为原时长的 1.9 倍）而目标短语节律不变（-sure or time）时，或语境部分节律不变而将目标词节律加快（时长为原本的 0.6 倍）时，受到语境节律的

chinaXiv:202010.00070v1

影响，被试更倾向于认为后缀“-sure”与功能词“or”之间无边界，而更少报告听到了“or”，同时改变语境与目标词的节律时，被试报告听到“or”的概率与正常语速无差异（类似结果见于 Maslowski et al., 2019）。Kösem 和 Wassenhove（2016）采用 MEG 记录方法在神经水平上进行了重复，发现语境节律不变、而目标词节律突然发生改变时，原有语境边界节律引发的“夹带作用”仍会持续一段时间，并对目标词的语音感知结果产生影响。这些结果提示，“夹带”作用不仅与语音声学包络的自下而上加工有关，还可能反映自上而下的、主动的时间预期对韵律边界感知的影响。

在此基础上，研究者试图从“夹带”这一“混合物”中分离出与韵律边界的自上而下加工机制有关的频段、时间进程与脑网络特点。Park et. al（2015）发现 delta 频段和 theta 频段神经振荡活动可能与韵律边界的自上而下加工有关。他们让被试聆听 7 分钟的故事语音，同时记录了 MEG 信号，基于因果性时间相关分析，以听倒转语音为基线条件，采用在一定时间延迟条件下其他脑区对听皮层脑区的神经信号相位的预测性作为该脑区对听皮层进行自上而下调控的神经指标（TDI, top-down index）。结果发现，锁时在韵律边界处（edge）进行分析发现，不同脑区对左侧听皮层的调控作用分别在边界前或边界后达到峰值，研究者推测这可能与各脑区在时间预期中扮演的角色有关。左侧额下回脑区可能通过 delta 频段振荡活动调控左侧听皮层活动，在实际声学信号输入前产生时间预期，而左侧运动前区和右侧额-颞联合区可能基于边界后已经输入的声学信号，通过 theta 频段振荡活动实时调控时间预期的匹配过程，进行时间预期模型的更新。遗憾的是，该研究没有对韵律边界的声学线索、韵律层级或其它自上而下因素进行操纵，无法为韵律边界的时间预期提供直接证据。Kayser et al.（2015）以故事语音为材料，操纵了语音之间“停顿”（pause）的变异性（jitter），变化率设置为 0%、30%、60%和 90%，而保持总体节律和可理解性不变。锁时在异常“停顿”与之后词汇起音（onset）的边界处进行分析发现，在边界前（-0.16～-0.06s）左前额脑区 alpha 频段（8~12Hz）能量降低，在边界后额-颞脑区 delta 频段（0.5~2Hz）神经振荡活动和语音包络之间的相位一致性降低，即对底层输入信息的“夹带”作用减弱，且二者存在显著正相关，其他频段对语音包络的“夹带”活动和 ERP 效应与正常条件相比无显著差异。这一结果表明，在语音边界的声学信号输入前，左侧前额脑区可能就已经产生了韵律边界时间预期（alpha 频段能量降低），并将预期信号传至听觉感知脑区，从而实时调控韵律边界的在线加工的过程，delta 频段神经振荡活动可能是大脑基于语音节律对韵律边界进行时间预期的产物，而与自下而上声学信号加工相对无关。以上研究结果揭示，前额叶脑区到颞叶脑区的感

chinaXiv:202010.00070v1

知-运动系统可能参与韵律边界的自上而下时间预期过程，并通过前额区 alpha 频段、额-颞区 delta 频段或 theta 频段神经振荡活动对语音流的在线切分起调控作用。

对于韵律边界的自上而下预期机制，研究者较少从理论角度进行深入阐释。然而，以上研究结果发现，听者自下而上的感知加工过程及其结果会受到自上而下预期加工的影响，这可能反映了一种“合成*分析”加工过程。“合成*分析”理论（Analysis-by-Synthesis Hypothesis: Halle & Stevens, 1962）认为，听者会基于对已获得的（声学或语言学）语境信息的分析，“合成”对即临声音信号模式的假设。在实际信号输入后，听者基于内部模型的“预期”信号与外部输入的信号进行“分析”，从而获得对长时程语流的实际感知结果。当二者匹配时，则能够促进感知加工过程。基于“合成*分析”理论，对韵律边界的自上而下加工与自下而上加工存在动态的交互过程。当一段语句或对话展开时，听者可能基于其长时程的韵律特点“合成”对韵律边界出现时间的预期模型，并基于预期模型对自下而上输入的韵律边界感知产生即时影响，如改变对连续语流的切分方式(e.g. Dilley & Pitt, 2010; Maslowski et al., 2019)、或提高语音感知质量（e.g. Park et al., 2015）。在语义、句法加工上，这种“合成*分析”方式已被证明能够提高口语加工效率，那么对韵律边界的加工是否也遵循“合成*分析”过程呢？对韵律边界的自上而下加工机制和自下而上加工机制的交互过程是如何的、是否存在时空特点上的分离？这些问题仍然有待研究。

4.3 CPS 与神经振荡的关系

CPS 和低频段的大脑神经振荡活动都是韵律边界感知的重要神经指标，近来的研究进一步提示二者之间可能存在一定的对应关系，即他们都对短语层级以上的韵律边界更为敏感。一些研究者发现，对于韵律边界加工引起的 CPS 效应，存在“长度限制效应”，即 CPS 效应在特定语音结构长度的边界处才会产生。Li 和 Yang（2009）采用中文句子为材料发现，在韵律词边界处未诱发显著的 CPS。类似地，Holzgrefe et al.（2013）利用并列名单（Pat or Jay and Lee）作为实验的语言材料，发现如果是将结构解析为（Pat）（or （Jay and Lee）），即早闭合结构，则 Pat 之后的边界并没有诱发 CPS， 如果将结构解析为（Pat or Jay） and Lee，则 Jay 之后诱发了 CPS。对于这一结果，作者认为 CPS 反映了句子理解过程中韵律边界信息的整合，也即，边界是否被用于句子理解才是诱发 CPS 的关键。然而，Pannekamp et al.（2005）证实 CPS 是独立于语言信息存在的，该说法与之冲突。因此，整合一说似乎不足以解释。另一种可能是一定长度的语音结构边界才能诱发 CPS，如韵律短语层级及以上（e.g. Li & Yang, 2009），这一解释与 Hwang & Steihauer （2011）所开展的隐性韵律研究所得结果

chinaXiv:202010.00070v1

一致。在大脑神经振荡活动中，韵律短语及以上层级对应于 delta 频段（0～3Hz），这一频段的“夹带”作用被证明在韵律边界的自下而上追踪(e.g. Doelling et al., 2014; Brodbeck et al., 2018)和自上而下预期（e.g. Kayser et al., 2015; Park et al., 2015）中扮演了重要角色。

鉴于 CPS 和 delta 频段神经振荡活动都与短语及以上韵律层级的边界感知有关，其背后反映的认知意义是如何的呢？结合自然语言的节奏特点及大脑对语言加工的特点，二者实际上可能反映了大脑以短语为组块（chunking）切分语音流的偏好。语料库分析结果显示，人类语言中语调短语（即一个语调韵律单位）是 4-8 个音节居多（周游，刘方舟, 2017），而这一节律恰好与 delta 频段节律相对应。根据神经认知研究结果，大脑对外部信息存在两个优势的整合时间窗口，分别为 150-300ms 的短时窗 （右偏化）和 3s 的长时窗，对应于 theta 频段（4～8Hz）、delta 频段（0～3Hz）的神经振荡活动（杨玉芳，2016），这些频段（尤其是 delta 频段）被证明和韵律边界的时间预期有关。一项固定步速阅读研究发现，人们在阅读时也会在短语边界产生 CPS，且 CPS 以 2～3s 为周期循环出现（Roll et al., 2012; Schremm et al., 2015），与 delta 频段节律一致。这些研究提示，基于日常口语中语音组块的节奏规律，大脑可能倾向于以一定的语音结构长度组块（如：短语）对语音流进行切分与整合加工，在这一组块单元中进一步对音节、词汇进行切分，或连接不同组块单元形成更高层级句子、段落结构的感知。CPS、delta 频段神经振荡活动与短语层级以上的韵律边界加工之间的密切关系，有可能反映了大脑的这种组块偏好（Meyer et al., 2017）。对于三者的关系，可以在未来研究中通过操纵语言材料本身的组块规律（采用不同节奏特点的语言）、或基于内隐学习可获得的组块规律（诱导大脑以某种规律切分语言）及其与大脑 CPS 反应、delta 频段振荡活动的关系来进一步验证。此外，这种组块偏好是否在语言理解过程中普遍存在、对语言理解的重要性是如何的，这些问题同样值得研究。

5 韵律边界加工能力的发展

韵律边界对言语理解至关要，韵律边界是言语理解中比句法结构解析更早利用的语言线索。然而，对韵律边界的感知却不是先天获得的，而是后天习得的结果且随年龄增长消退，同时还受其它因素影响。

5.1 母语者韵律边界加工能力的发展

出生不久的婴儿已经可通过韵律边界提取句法信息。虽然婴儿脑中对韵律边界的加工尚处于较低水平，但是婴儿仍然能够通过感知韵律边界分割语音流（Hawthorn & Gerken, 2014）。

chinaXiv:202010.00070v1

然而，不同年龄段的孩子，对声学线索的利用不同。4 个月大的英语婴儿，在听辨从句边界的时候，须依靠全部的三个声学线索（Seidl & Cristia, 2008），而 6 个月大的英语婴儿则主要依赖音高（Seidl, 2007），在另一个以 3 岁和 6 岁儿童为研究对象的研究中，研究者发现 3 岁儿童对停顿线索的依赖优先于音长线索，而必需要到 6 岁才可以不依靠停顿，仅仅借助音长和音高来判断韵律边界（Holzgrefe-Lang et al., 2016; Männel et al., 2013）。也就是说，对于 6 岁以上的孩子来说，音高和音长的组合就可以实现韵律边界的感知，引发 CPS。显然，随着年龄的增长，儿童对边界声学线索的倚重产生了变化。Ommena et al.（2020）从对比德语母语儿童和法语母语儿童（6 个月和 8 个月）在听辨法语并列名单的表现中发现，由于德语的韵律边界比法语边界更倚赖界前延长和音高，8 个月的德语母语儿童比 6 个月的儿童发展出了对界前延长和音高的敏感性，而法语儿童则没有出现这种现象。该结果显示声学线索权重的发展变化是受母语影响的。

以上现象被称为发展性权重转移。这种现象在语音习得和韵律习得中都有发生（Seidl, 2007）。Seidl 和 Crestia（2008）发现，儿童必须经历一个发展性权重转移的过程才能获得和成人一样的韵律边界声学线索利用策略，且转移的路径是从整体（holistic）到分析（analytic）。已有研究提示语言的熟练水平影响儿童韵律边界感知能力的发展，但需要注意的是，儿童发展语言的同期也存在认知水平的发展。那么，认知发展在韵律边界感知能力的发展中是否也扮演一定的角色有待进一步研究。

另一方面，Steinhauer et al. （2010） 发现虽然韵律边界线索对老年人的句法分析也具有引导作用，帮助他们解决句法歧义，但对比青年人，老年人的加工方式与年青人基本一致，对句法和韵律边界的冲突却不如年轻人敏感。该研究结果似乎表明，与婴幼儿的韵律边界感知能力随其年龄迅速发展相反的是，老年人利用韵律边界线索的能力随年龄的增长而下降。

综上所述，韵律边界的感知随着年龄的增长而发展，但是到达一定年龄后，它有可能随着年龄的增长而退化。这似乎暗示韵律边界的感知能力和认知能力的发展有相关关系。

5.2 二语者韵律边界加工能力的发展

韵律边界提供关于句子结构和句子意思的关键信息，是母语者在言语加工中所依赖的重要韵律线索（Nickels et al., 2013）。二语学习者是否像母语者一样加工目标语言的韵律边界呢？对此，学界有不同的观点。

Nickels et.al.（2013）以英语母语者和德-英（L2）学习者为被试，以停顿为韵律边界（语调边界）的操控手段，以自然度判断为实验任务考察不同语言背景的听话人对韵律边

chinaXiv:202010.00070v1

界的敏感性。如（1）When a bear is approaching the people#1 the dogs come running. （2）When a bear is approaching # 2 the people#3 come running.这样的口语句子，四个条件分别为句中 a.无任何语调韵律边界; b.#1 处有边界;c.#2 处有边界; d. #2 和#3 处都有边界。 结果发现二语学习者在判断自然度的得分上与母语者没有显著差异，脑电成分仅在时间上比母语者稍有延迟，但总体模式是一样的。在其进一步的研究（2018）中，被试群体扩展为母语者和不同语言距离的两个学习者群体，德-英（L2）学习者和汉-英（L2）学习者，其他实验组成不变。结果发现，汉语学习者也反应出与母语者和德语学习者趋势一致的行为和脑电结果，但是条件 d 情况下的接受度显著高于德语和英语者，反映错误边界（即#3 处）的脑电幅度并没有像德语学习者一样随目标语言熟练度增大而增大，似乎更倾向于自下至上的加工方式，而不是另两个组别所显示的自上而下的预测性的加工方式。对于汉语学习者的表现，作者解释为也许和所采用的动词有关，也许和边界前韵律廓形的知觉有关。以上两个实验似乎暗示，韵律边界的应用可能涉及到的不仅仅是边界声学线索的知觉，而且还涉及到多个维度甚至是多个模态线索的整合，并且该过程受母语经验的影响。未来的研究应该以全局观来看待韵律边界的感知问题。

但也有研究者持有不同的观点，他们认为二语者不能像母语者一样运用韵律边界。Pennington 和 Ellis （2000） 选择粤-英（L2）双语被试，所采用的任务是辨别当下听到的英语句子是否与先前听过的句子一致。听辨任务分成两个阶段。第一个阶段，被试听 24 个目标句子。第二阶段，再听 48 个测试句，并判断测试句中有哪些是第一阶段出现过的。测试句分成三类，第一类是词语不同但句法相同的句子，第二类是从目标句中挑选出来的句子，第三类是词语相同但是韵律线索不同的句子。结果发现，被试在判断第一类和第二类句子的表现远好于随机水平，但是第三类句子却远低于随机水平。即便是被提醒注意韵律方面的差异（第二个实验），结果没有改善。对此，作者解释为，在课堂学习的外语者，尤其是处在具有背诵文化的教育环境下，是习惯于将韵律特征和句子意义捆绑在一起学习的，没有专门对语音线索的感知进行培训。除此之外，还有可能是学习者并不懂得韵律线索（包括边界）和句法/句意之间的投射关系。学习者的失败究竟是因为未能感知到边界声学线索还是不能进行正确的投射，又或是由于任务对学习者的认知资源消耗太多，使得学习者无暇分心关注韵律线索？然而，作者并没有进一步探查。Schmidt et al.（2019）的观点来源于记忆实验。他们以英语母语者和希-英（L2）学习者为被试，以数字记忆（测量的方式类似于数字广度记忆测量法）为任务，比较有韵律边界和无韵律边界的条件下，母语者和外语学习者对数字

chinaXiv:202010.00070v1

的记忆成绩。结果发现，在有韵律边界的情况下，两个群体对母语数字的记忆成绩都好于没有韵律边界下的记忆，但外语数字的记忆成绩则不受韵律边界存留的影响。对于该现象，作者认为韵律边界是深层结构，因而外语者不能像母语者一样应用。该研究的结论与传统观点不一致。传统上，韵律边界是语言的表层结构。无疑，这个观点为韵律边界的认识带来了新的方向。以上的四个研究给二语韵律边界感知发展领域留下了多个值得进一步探索的问题。

训练可以促进韵律边界的感知能力。Yang （2016）在课堂教学中发现通过跟读（shadowing）目标语言的表达方式或模仿的方式再现自然声音模式，可以帮助学习者在头脑中再现说话者韵律边界和句法理解的联系。目前已有的结果并没有揭示训练影响韵律边界感知的机制，我们不清楚是什么因素影响了边界感知，也不清楚边界感知中哪个具体的线索被影响。鉴于外语学习者普遍反映的听力困难、母语者说得太快等现象，也可能是由于韵律边界感知缺失造成的，韵律边界的感知习得机制和训练值得更深入的研究。

6 韵律边界加工的未来展望

韵律边界是韵律中的一个物理表征，与重音和语调相比，具有更多的认知功能，是我们观察口头语言加工和习得的重要窗口，因而被众多研究作为观察对象和操控对象。通过对心理学和语言学领域的韵律边界研究进行全面的总结和分析，可以发现，关于韵律边界的研究还有很多问题需要进一步的探讨。

6.1 韵律边界的声学表现需要进一步探讨

韵律边界语音声学信号加工的研究结果非常丰富，研究者们基于母语者的加工，对韵律边界的结构等级、强度的声学线索表征、韵律边界的感知权重分配(包括单线索和线索组合）等方面做了深入的探查。这些研究结果为听力障碍辅助器材的设计（如人工耳蜗）、语音识别和人工语言的合成提供了坚实的实证基础。然而，已有研究未能将句法的影响一同考查，考虑到韵律边界和语言加工的紧密关系，这方面的研究还需要进一步加强。另外，这类研究绝大部分在实验室开展的，其使用的实验材料以朗读素材（reading speach）为主，因而削弱了生态效度。未来的研究可以扩大研究面，以不同文体风格的口头语言为素材，比如讲故事、朋友间的随性聊天等，在不同的背景环境下（比如有/无噪音）考查听者对韵律边界声学线索特征的运用情况。最后，未来对待声学线索感知特征的描述可以进一步细化，不要仅仅考查单个线索的存留对边界感知的影响，而是以连续性的视角（例如，阈值）来看待声学线索对感知的影响。

6.2 韵律边界的加工过程需要进一步明确

chinaXiv:202010.00070v1

以往研究者主要通过 CPS 和大脑低频段的“夹带”作用（即神经振荡活动）阐述了韵律边界的加工机制，证实了韵律边界能够独立于其他语言信息被独立加工、并且和语言结构的感知、语义或句法的理解密切相关。但是，对于韵律边界的认知神经机制，仍然存在许多未解决的问题。首先，关于韵律边界的认知神经机制的时空特点仍有许多问题需要进一步考察。例如，以往研究初步发现，“夹带”作用能够分别反映听觉脑区对韵律边界的自下而上加工过程、或高级额叶脑区基于对韵律边界的预期对颞-顶区域的自上而下调控过程，但对这两种过程产生交互作用时各自发挥作用的具体时间进程和脑网络特点缺乏系统的讨论。此外，不同于其他语言线索，声学韵律边界并非与语义边界、句法边界存在一一对应关系，尚不清楚当预期韵律和底层输入的声学韵律产生冲突时，哪一过程会以更高的相对权重影响韵律边界感知。再者，近来研究提示，CPS、delta 频段的“夹带”作用和大脑偏好以短语为单位对语言进行组块，三者的关系需要进一步验证，在二语者、特殊人群中拓展研究也将有利于理解大脑对韵律边界的加工特点和语言特点、发展的关系。第二，尚不清楚对韵律边界的“夹带”作用的认知重要性是如何的。在安静环境中进行口语理解时，韵律被认为是一种辅助而非必要的线索，即使在去除韵律音高波动信息（Xu et al., 2013）、或去除停顿线索（e.g. Li & Yang, 2009; Kayser et al., 2015）时，听者仍然能够正常理解口语句子，而在噪声场景中对韵律结构的感知能力和口语感知成绩存在相关（Slater & Kraus, 2016）。对韵律边界的自上而下时间预期是一种韵律边界加工过程中普遍应用的机制（即不依赖于场景条件而存在），还是只在困难条件下（如噪声掩蔽、语音音质受损等）才会发挥作用？这些问题有待未来研究讨论。最后，许多研究证据表明韵律边界加工会影响高级语言加工（例如语义、句法加工），在个体水平上，对语言韵律时间结构的加工能力和其进行语音感知、阅读理解的能力有关（e.g. Cason et al., 2015; Cumming et al., 2015; Flaugnacco et al., 2015; Hidalgo et al., 2017; Schön & Tillmann, 2015; Slater & Kraus, 2016），这说明不同个体基于韵律边界加工促进高级语言功能的具体机制可能存在差异，但目前的研究较少直接讨论这一点。探究这一问题将有助于理解韵律边界加工和其他高级语言认知功能的潜在关系、以及为一些语言能力发展不足（例如阅读障碍儿童）或退化的人群（例如老年人、存在语言障碍的患者）提供可能的干预方案。

6.3 韵律边界加工与句法加工的关系需要进一步厘清

已有研究发现韵律边界影响句子加工，句法可以预测韵律边界，而且二者相互制约。这些观点得到了行为实验和神经科学证据的支持。但是，这些研究结果是基于语义、句法和韵律是分层的假设（Ischebeck et al., 2008），其研究框架主要是韵律信息（和其它非句法信息，

chinaXiv:202010.00070v1

如词汇、语用等）和句法信息在句子加工中的地位和关系。即，听一句话的时候是首先使用句法信息对句子进行分解还是同时运用多个信息对句子进行解析。其中，韵律边界线索处于辅助地位。然而，已有的结果显示韵律边界和句法的关系并不是一对一的简单关系，而是要比我们知道的复杂得多。韵律边界可能不仅仅是一个通过和句法相互作用而参与句子理解的表层结构，其在本质上可能就是语言规则的具身表现（embodiment），是将抽象的规则、物理世界和人类体验联结在一起的接口（Giraud & Poeppel, 2012; Kreiner & Eviatar, 2014）。韵律边界的大小（边界的大小以声学特征来表征）表现了句子成分之间关系的紧密程度，再现言语的层级组织结构，以体现语言信息在时域上布局的方式，方便大脑通过振荡模拟其节奏，从而捕捉到抽象的语言组织规则。比如，韵律边界长度的研究发现韵律边界特异性脑电成分 CPS 并不能仅仅被声学线索诱发，而是对一定长度组块（如：短语）对语音流进行切分与整合加工的反应。又如 Ding et.al. （2016） 的跨语言研究发现在次语调韵律边界所产生的 EEG 成分的强度显著地大于韵律单位内部位置上产生的强度（见前文）。未来的研究可以跳出传统的结构语言学框架，而从具身研究的角度，利用 EEG 和 MEG 技术，使用不同节奏规律的语言、滤波和语音反转等手段操纵韵律边界的感知，检视韵律边界感知的神经振荡特征和句法解析之间的关联性，从而深入揭示韵律边界的实质和功能，进一步厘清韵律边界加工与句法加工的关系。

6.4 韵律边界加工能力的发展需要进一步关注

韵律边界的声学线索加工能力与语言理解息息相关，但我们对二语者的韵律边界加工能力发展情况所知不多。遗憾的是，以二语者加工韵律边界为主题的文献极少。我们不清楚二语者与母语者韵律边界的加工差异到底存在于哪些具体的方面，比如是对具体某个声学线索的知觉差异吗？也不清楚导致这些差异的因素。是二语者采用不同于目标语的线索感知权重模式引起的不同？还是二语者所采用的加工机制不同，比如采用不同于母语者的加工方向而引起不同？虽然已有研究提出背景语言差距大的二语者可能更倾向于使用自下而上的线索（Nickels & Steinhauer, 2018），但并没有研究专门回答该问题。除此之外，也有人提出了 CPS 实际上是整合结果的体现（Nickels, 2013），那么，二语者是否也具备和母语者一样的整合能力呢？从神经学层面来看，二语者和母语者加工韵律边界的时间进程和脑网络是否存在不同的特点？这些问题都值得进一步探究。

听力理解困难一直是二语教学的重大问题。外语学习者常常抱怨母语者说话太快，其背后原因可能是二语者不能有效地利用韵律边界将语流进行切分。如果能对韵律边界的感知能

chinaXiv:202010.00070v1

力进行培训的话，那么二语听力困难将有一个全新的解决方案。然而，目前针对二语者韵律边界加工能力发展过程的研究也几乎为零。二语者的边界感知能力是如同母语者一样从整体到局部地发展，还是受母语迁移的影响而产生石化（fossilizaiton）现象？又或者是逐步发展出符合目标语加工的次级技能（subskills），最终接近目标语母语者的水平？如果是这样，那么韵律中介语的发展路径又是如何,是否受母语经验的影响而具有特异性的发展途径呢？教学干预是否能够提升韵律边界加工能力？例如利用边界声学线索感知训练增强自下而上的加工能力或者利用节奏训练增加其自上而下的加工能力？要回答上述问题,未来的研究必须关注学习者二语韵律边界声学线索的在线加工机制、习得过程和不同训练方法的效益问题。除此之外，还应该关注个体在声学线索感知上的差异可能带来的影响，以便在教材编写和教学设计上能够更有针对性。

chinaXiv:202010.00070v1

参 考 文 献

杨玉芳. (2016). *心理语言学*. 北京：科学出版社.

周游,刘方舟. (2017). 汉语韵律结构与语法结构的对比分析. 见中国中文信息学会语音信息专业委员会(编)，*第十四届全国人机语音通讯学术会议(NCMMSC'2017)论文集* (pp.205–210).中国中文信息学会语音信息专业委员会清华信息科学与技术国家实验室.

Aasland, W., & Baum, S. (2003). Temporal parameters as cues to phrasal boundaries: A comparison of processing by left–hemisphere–damaged and right–hemisphere–damaged individuals. *Brain and Language*, *87*(3), 385–399.

Bögels, S., Schriefers, H., Vonk, W., Chwilla, D. J., & Kerkhofs, R. (2010). The interplay between prosody and syntax in sentence processing: The case of subject-and object-control verbs. *Journal of Cognitive Neuroscience*, *22*(5), 1036–1053.

Bögels, S., Schriefers, H., Vonk,W., Chwilla, D. J. & Kerkhofs, R. (2013). Processing consequences of superfluous and missing prosodic breaks in auditory sentence comprehension. *Neuropsychologia*, *51*(13), 2715–2728.

Brodbeck, C., Hong, L. E., & Simon, J. Z. (2018). Rapid transformation from auditory to linguistic representations of continuous speech. *Current Biology*, *28*(24), 3976–3983.

Buxó-Lugo, A. & Watson, D.G. (2016). Evidence for the influence of syntax on prosodic parsing. *Journal of Memory and Language*, *90*, 1–13.

Clahsen, H., & Felser, C. (2018). Some notes on the shallow structure hypothesis. *Studies in Second Language Acquisition*, *40*(3), 693–706.

Cason, N., Astésano, C., & Schön, D. (2015). Bridging music and speech rhythm: Rhythmic priming and audio–motor training affect speech perception. *Acta Psychologica*, *155,* 43–50.

Clifton,C. & Duffy,S. (2001). Sentence and text Comprehension: Roles of linguistic structure. *Annual Review of Psychology*, *52*,167–196.

Clifton, C., Carlson, K., & Frazier, L. (2002). Informative prosodic boundaries. *Language and Speech*, *45*(2), 87–114.

Cole, J., Mo,Y & Baek,S. (2010). The role of syntactic structure in guiding prosody perception with ordinary listeners and everyday speech. *Language and Cognitive Process*, *25*(7), 1141–1177.

Cole, J. (2015). Prosody in context: A review. *Language, Cognition and Nuroscience*, *30*(1),1–31.

Cumming, R., Wilson, A., Leong, V., Colling, L.J., & Goswami, U. (2015). Awareness of rhythm patterns in speech and music in children with specific language impairments. *Frontiers in Human Neuroscience*, *9*, 1–21.

Dilley, L. C., & Pitt, M. A. (2010). Altering context speech rate can cause words to appear or

chinaXiv:202010.00070v1

chinaXiv:202010.00070v1

disappear. *Psychological Science*, *21*(11), 1664–1670.

Ding, N., Chatterjee, M., & Simon, J. Z. (2014). Robust cortical entrainment to the speech envelope relies on the spectro-temporal fine structure. *Neuroimage*, *88*, 41–46.

Ding, N., Melloni, L., Zhang, H., Tian, X., & Poeppel, D. (2016). Cortical tracking of hierarchical linguistic structures in connected speech. *Nature Neuroscience*, *19*, 158-164.

Ding, N., Pan, X., Luo, C., Su, N., Zhang, W., & Zhang, J. (2018). Attention is required for knowledge-based sequential grouping: Insights from the integration of syllables into words. *Journal of Neuroscience*, *38*(5), 1178–1188.

Ding, N., & Jin, P. (2019). Low–frequency neural activity reflects rule-based chunking during speech listening. bioRxiv, Article 742585v1. https://www.biorxiv.org/content/10.1101/742585v1.full.pdf

Doelling, K. B., Arnal, L. H., Ghitza, O., & Poeppel, D. (2014). Acoustic landmarks drive delta–theta oscillations to enable speech comprehension by facilitating perceptual parsing. *Neuroimage*, *85*, 761–768.

Flaugnacco, E., Lopez, L., Terribili, C., Montico, M., Zoia, S., Schön, D. (2015). Music training increases phonological awareness and reading skills in developmental dyslexia: A randomized control trial. *PLOS ONE*, *10*(9), Article e0138715. https://doi.org/10.1371/journal.pone.0138715

Frazier, L., Clifton, C., & Carlson, K. (2004). Don't break, or do: Prosodic boundary preferences. *Lingua*, *114* (1), 3–27.

Frazier, L., Carlson, K., & Clifton, C. (2006). Prosodic phrasing is central to language comprehension. *Trends in Cognitive Sciences*, *10*(6), 244–249.

Giraud, A. L., & Poeppel, D. (2012). Cortical oscillations and speech processing: Emerging computational principles and operations. *Nature Neuroscience*, *15*(4), 511–517.

Gross, J., Hoogenboom, N., Thut, G., Schyns, P., Panzeri, S., Belin, P., & Garrod, S. (2013). Speech rhythms and multiplexed oscillatory sensory coding in the human brain. *PLOS Biology*, *11*(12). Article e1001752. https://doi.org/10.1371/journal.pbio.1001752

Halle, M., & Stevens, K. (1962). Speech recognition: A model and a program for research. *IRE Transactions on Information Theory*, *8*(2), 155-159.

Hawthorn, K. & Gerken, L. (2014). From pauses to clauses: Prosody facilitates learning of syntactic constituency. *Cognition, 133*(2), 420–428.

Hidalgo, C., Falk, S., & Schön, D. (2017). Speak on time! Effects of a musical rhythmic training on children with

hearing loss. *Hearing Research*, *351*, 11–18.

Holliman, A.J., Gutiérrez Palma, N., Critten, S., Wood, C., Cunnane, H., Pillinger, C. (2017). Examining the independent contribution of prosodic sensitivity to word reading and spelling in early readers. *Reading and Writing*,*30*, 509–521.

Holzgrefe, J., Wellmann, C., Petrone, C., Truckenbrodt, H., Höhle, B., & Wartenburger, I. (2013). Brain response to prosodic boundary cues depends on boundary position. *Frontiers in Psychology Language Sciences*, *4*, 421.

Holzgrefe-Lang. J., Wellmann, C., Petrone, C., & Räling, R. (2016). How pitch change and final lengthening cue boundary perception in German: Converging evidence from ERPs and prosodic judgments. *Language, Cognition and Neuroscience*, *31*(7), 1–17.

Honbolygo, F., Török, Á., Bánréti, Z., Hunyadi, L., & Csépe, V. (2016). ERP correlates of prosody and syntax interaction in case of embedded sentences. *Journal of Neurolinguistics*, *37,* 22–33.

Howard, M. F., & Poeppel, D. (2010). Discrimination of speech stimuli based on neuronal response phase patterns depends on acoustics but not comprehension. *Journal of Neurophysiology*, *104*(5), 2500–2511.

Hwang, H., & Steinhauer, K. (2011). Phrase length matters: The interplay between implicit prosody and syntax in korean "garden path" sentences. *Journal of Cognitive Neuroscience*, *23*(11), 3555–3575.

Ischebeck, A. K., Friederici, A. D., & Alter, K. (2008). Processing prosodic boundaries in natural and hummed speech: An fMRI study. *Cerebral Cortex*, *18*(3), 541–552.

Kayser, S. J., Ince, R. A.A., Gross, J., & Kayser, C. (2015). Irregular speech rate dissociates auditory cortical entrainment, evoked responses, and frontal alpha. *Journal of Neuroscience*, *35*(44), 14691–14701.

Kjelgaard, M. M., & Speer, S. R. (1999). Prosodic facilitation and interference in the resolution of temporary syntactic ambiguity. *Journal of Memory and Language*, *40*(2), 153–194.

Kösem, A., & Wassenhove, V. (2016). Distinct contributions of low- and high-frequency neural oscillations to speech comprehension. *Language,Cognition and Neuroscience*, *32*, 536–544.

Kreiner, H., &Eviatar, Z. (2014). The missing link in the embodiment of syntax: Prosody. *Brain and Language, 137,* 91-102.

Krivokapi, J., & Byrd, D. (2012). Prosodic boundary strength: An articulatory and perceptual study. *Journal of Phonetics*, *40*(3), 430–442.

Li, W. J., & Yang, Y. (2009). Perception of prosodic hierarchical boundaries in Mandarin Chinese sentences. *Neuroscience*, *158*(4), 1416–1425.

Männel, C., Schipke, C. S., & Friederici, A. D. (2013). The role of pause as a prosodic boundary marker: Language

chinaXiv:202010.00070v1

ERP studies in German 3– and 6–year–olds. *Developmental Cognitive Neuroscience*, *5*, 86–94.

Männel, C., & Friederici, A. D. (2016). Neural correlates of prosodic boundary perception in German preschoolers: If pause is present, pitch can go. *Brain Research*, *1632*, 27–33.

Marcus, M., & Hindle, D. (1990). "Description theory and intonation boundaries," In G. Altman (Eds.), Cognitive Models of Speech Processing: Psycholinguistics and Computational Perspectives (pp. 483–512). Cambridge, MA, USA: MIT Press.

Maslowski, M., Meyer, A. S., & Bosker, H. R. (2019). Listeners normalize speech for contextual speech rate even without an explicit recognition task. *The Journal of the Acoustical Society of America*, *146*(1), 179–188.

Meyer, L., Henry, M. J., Gaston, P., Schmuck, N., & Friederici, A. D. (2017). Linguistic bias modulates interpretation of speech via neural delta-band oscillations. *Cerebral Cortex*, *27*(9), 4293–4302.

Nickels, S., Opitz, B., & Steinhauer, K. (2013). ERPs show that classroom-instructed late second language learners rely on the same prosodic cues in syntactic parsing as native speakers. *Neuroscience Letters, 557*, 107–111.

Nickels, S., & Steinhauer, K. (2018). Prosody–syntax integration in a second language: Contrasting event-related potentials from German and Chinese learners of English using linear mixed effect models. *Second Language Research*, *34*(1), 9-37.

Ommena, S.van., Boll-Avetisyanb,N.,Larrazaa,S.,Wellmannb,C.,Bijeljac-Babica,R., Höhleb,B., & Nazzi,T. (2020). Language-specific prosodic acquisition: A comparison of phrase boundary perception by French- and German-learning infants. *Journal of Memory and Language*,*112*,104-108.

Pannekamp, A., Toepel, U., Alter, K., Hahne, A., & Friederici, A. D. (2005). Prosody–driven sentence processing: An event-related brain potential study. *Journal of Cognitive Neuroscience*, *17*(3), 407–421.

Park, H., Ince, R. A., Schyns, P. G., Thut, G., & Gross, J. (2015). Frontal top-down signals increase coupling of auditory low-frequency oscillations to continuous speech in human listeners. *Current Biology*, *25*(12), 1649–1653.

Pauker, E., Itzhak, I., Baum, S. R., & K. Steinhauer. (2011). Effects of Cooperating and Conflicting Prosody in Spoken English Garden Path Sentences: ERP Evidence for the Boundary Deletion Hypothesis. *Journal of Cognitive Neuroscience*, *23*(10), 2731–2751.

Pennington, M., & Ellis, N.C. (2000). Cantonese Speakers' Memory for English Sentences with Prosodic Cues. *The Modern Language Journal*, *84*(3), 372–389.

Petrone, C., Truckenbrodt, H., Wellmann, C., Holzgrefe-Lang, J., Wartenbruger, I., & Hohle, B.(2017). Prosodic boundary cues in German: Evidence from the production and perception of bracketed lists. *Journal of Phonetics*,

chinaXiv:202010.00070v1

61,71–92.

Poeppel, D. (2003). The analysis of speech in different temporal integration windows: Cerebral lateralization as 'asymmetric sampling in time'. *Speech Communication, 41*(1), 245–255.

Qin, Z., Chine, Y., & Tremalay, A. (2016). Processing of word-level stress by Mandarin-speaking second language learners of English. *Applied Psycholinguistics*, *38*(3), 1–30.

Roll, M., Lindgren, M., Alter, K., & Horne, M. (2012). Time-driven effects on parsing during reading. *Brain and Language*, *121*(3), 267–272.

Schmidt, E., Pérez, A., Cilibrasi L., & Tsimpli, I. (2019). Prosody facilitates memory recall in L1 but not L2. *Studies in Second Language Acquisition*, *42*(1), 223–238.

Schön, D., & Tillmann, B. (2015). Short- and long-term rhythmic interventions: Perspectives for language rehabilitation. *Annals of the New York Academy of Sciences, 1337*, 32–39.

Schremm, A., Horne, M., & Roll, M. (2015). Brain responses to syntax constrained by time-driven implicit prosodic phrases. *Journal of Neurolinguistics*, *35,* 68–84.

Seidl, A. (2007). Infants' use and weighting of prosodic cues in clause segmentation. *Journal of Memory & Language*, *57*(1), 24–48.

Seidl, A., & Cristià, A. (2008). Developmental changes in the weighting of prosodic cues. *Developmental Science*, *11*(4), 596–606.

Selkirk, E. (2005). Comments on intonational phrasing in English, In S. Frota, M. Vigario, and M. J. Freitas (Eds.), *Prosodies* (pp. 11–58). Berlin, Germany: Mouton de Gruyter.

Slater, J., & Kraus, N. (2016). The role of rhythm in perceiving speech in noise: A comparison of percussionists, vocalists and non-musicians. *Cognitive Processing*, 17(1), 79–87.

Steinhauer, K., Alter, K., & Friederici, A. D. (1999). Brain potentials indicate immediate use of prosodic cues in natural speech processing. *Nature Neuroscience*, *2*(2), 191–196.

Steinhauer, K., Shani, H. A., Pauker, E., Itzhak, I., & Baum, S. R. (2010). Prosody-syntax interactions in aging: Event-related potentials reveal dissociations between on-line and off-line measures. *Neuroscience Letters, 472* (2), 133–138.

Truckenbrodt, H. (1995). *Phonological phrases: Their relation to syntax, focus and prominence* (Unpublished doctorial Dissertation), Massachusetts Institute of Technology.

Wagner,M., & Watson, D.G. (2010). Experimental and theoretical advances in prosody: A review. *Language and*

chinaXiv:202010.00070v1

Cognitive Process, *25*(7–9), 905–945.

Webman-Shafran, R., & Fodor, J. D. (2016). Phrase Length and Prosody in On-Line Ambiguity Resolution. *Journal of Psycholinguist Research,45*(3), 447–474.

Xu, G., Zhang, L., Shu, H., Wang, X., & Li, P. (2013). Access to lexical meaning in pitch–flattened Chinese sentences: An fMRI study. *Neuropsychologia, 51*(3), 550–556.

Yang, L. C. (2016, March). *Optimizing pronunciation and prosody teaching in second language learning*. Paper presented at the meeting of ISAPh 2016 International Symposium on Applied Phonetics, Nagoya, Japan.

Yang, X., Shen, X., Li, W., & Yang, Y. (2014). How listeners weight acoustic cues to intonational phrase boundaries. *PLOS ONE, 9*(7), Article e102166. http://doi:10.1371/journal.pone.0102166

Zoefel, B., & VanRullen, R. (2015). Selective perceptual phase entrainment to speech rhythm in the absence of spectral energy fluctuations. *Journal of Neuroscience, 35*(5), 1954–1964.

Zoefel, B., & VanRullen, R. (2016). EEG oscillations entrain their phase to high-level features of speech sound. *Neuroimage, 124,* 16–23.

Prosodic Boundaries in Speech: A Window to Spoken Language Comprehension

FANG Lan[1,2] ; ZHENG Yuanyi[3,4] ; JIN Han[2] ; LI Xiaoqing[3,4] ; YANG Yufang[3,4] ; WANG Ruiming[1]

（[1] Key Laboratory of Brain, Cognition and Education Sciences, Ministry of Education, & Center for Studies of Psychological Application, School of Psychology, South China Normal University, Guangzhou, 510631)

([2] School of Foreign Studies, Guangzhou University，Guangzhou 510030) ([3] Key Laboratory of Behavioral Science, Institute of Psychology, Chinese Academy of Sciences, Beijing 100101) ([4] Department of Psychology, Universities of Chinese Academy of Sciences, Beijing 100149)

Abstract: The perception of prosodic boundaries is critical to spoken language comprehension, and it has become a primary research topic among psychologists and psycholinguists in the past decade. Utterances are chunked into prosodic units of different strengths. The boundaries between prosodic units are mainly signaled by acoustic cues like pitch change, final lengthening, and pause. Previous cognitive, linguistic, developmental, and neuroimaging studies have significantly advanced our understanding of the processing of prosodic boundaries. We now know that listeners use a perceptual weighting strategy to process prosodic boundary cues, and there are specific brain mechanisms for

chinaXiv:202010.00070v1

prosodic boundary processing. The ability to processing prosodic boundaries steadily develops with age in young children and transfers to a second language, but it generally decreases with age in older adults. Future studies should expand the investigation of prosodic boundaries to more pragmatic genres and focus on revealing the cognitive mechanisms underlying prosodic boundary processing, the relationship between prosodic boundary and syntax processing, and the development of prosodic boundary perception in second language learners.

Key words：prosodic boundary; spoken language comprehension; acoustic characteristics; neural oscillation; development

chinaXiv:202010.00070v1